

Lógica y Filosofía de la Ciencia Máster Interuniversitario

Declaración de integridad intelectual

Trabajo Fin Máster

Curso 2016-2017

Título del trabajo: **Una evaluación de la postura expresivista de Dorit Bar-On con respecto al autoconocimiento**

1. Sé que copiar es una forma de deshonestidad académica.
2. He leído el documento sobre cómo ser intelectualmente íntegro, estoy familiarizado con sus contenidos y he evitado todas las formas de plagio allí recogidas.
3. Cuando utilizo las palabras de otros, lo indico mediante el uso de comillas.
4. He referenciado todas las citas e igualmente el resto de ideas tomadas de otros.
5. No he plagiado mi propio trabajo.
6. No permitiré a otros que plagien mi trabajo.

Fecha: **30/07/2017**



Firma: **Andrés Luna Bermejo**

Una evaluación de la postura expresivista de Dorit Bar-On con respecto al autoconocimiento

Andrés Luna Bermejo

Mastér Interuniversitario en Lógica y Filosofía de la Ciencia

Universidad de Salamanca

RESUMEN

Este trabajo presenta una crítica a la explicación neoexpresivista del autoconocimiento ofrecida por Bar-On. Bar-On defiende que nuestras autoadscripciones mentales son expresiones de nuestros estados mentales presentes, lo cual explicaría la especial seguridad que estas acarrearán; pero también sostiene que las susodichas autoadscripciones articulan conocimiento privilegiado de nuestros propios estados mentales presentes. El objetivo principal de este trabajo es negar que nuestras autoadscripciones mentales puedan ser consideradas afirmaciones de conocimiento privilegiado. Para ello, argumentaré en primer lugar que el estatus epistémico de nuestras afirmaciones se debe a su carácter meritorio, el cual está íntimamente ligado con la capacidad del sujeto para ofrecer razones a favor de sus afirmaciones. En segundo lugar, mostraré que, en tanto que expresiones, nuestras autoadscripciones carecen de dimensión epistémica, esto es, que no tiene sentido evaluarlas atendiendo al mérito o demérito.

Palabras clave: seguridad especial, perspectiva de la primera persona, autoadscripciones mentales, mérito, afirmaciones de conocimiento.

ABSTRACT

This paper is a critic of Bar-On's neo-expressivist account of self-knowledge. Bar-On argues that avowals are expressions of our present mental states, which would explain their special security; but Bar-On

also contends that avowals articulate privileged self-knowledge about our own present mental states. The main goal of this paper is to deny that avowals can be viewed as privileged knowledge claims. In order to that, firstly, I shall argue that the epistemic status of our claims depends on its meritorious nature, which is closely connected with the subject's capacity for providing reasons to support her claims. Then, I shall show that, as expressions, avowals do not belong to the epistemic realm, so it makes no sense to assess them with respect to their merit.

Key words: authority, first-person, avowals, merit, knowledge claims.

1. INTRODUCCIÓN

En este trabajo, nos centraremos en ofrecer una evaluación de la explicación del fenómeno del autoconocimiento ofrecida por Dorit Bar-On. Nuestro propósito central es negar que autoadscripciones mentales como “Me duele”, “Creo que...”, etc. llevadas a cabo desde la perspectiva de la primera persona sean afirmaciones de conocimiento. Bar-On defiende que ese tipo de proferencias constituyen expresiones de nuestra vida mental, esto es, exhiben, manifiestan, la condición psicológica del sujeto, de modo que cualquier observador competente podría tener acceso directo a su vida mental simplemente con comprender dicha proferencia. La presencia de una dimensión expresiva explicaría, según Bar-On, por qué nuestras autoadscripciones mentales poseen una seguridad especial o autoridad de la primera persona. Esta respuesta no-epistémica al fenómeno de la autoridad de la primera persona es conjugada por Bar-On con una explicación de la dimensión epistémica de las autoadscripciones que hace uso de las nociones de creencia, verdad y justificación. Nuestra principal objeción a Bar-On será que su explicación no da cuenta de la dimensión epistémica de nuestras autoadscripciones mentales, pues el aspecto definitorio de las afirmaciones de

conocimiento no es otro que su carácter meritorio, del cual carecen nuestras autoadscripciones mentales en perspectiva de primera persona.

De acuerdo con este fin, dividiremos la argumentación en tres secciones: en primer lugar, presentaremos la cuestión del autoconocimiento: qué es el autoconocimiento, qué perspectiva ocupa el sujeto que se conoce a sí mismo, etc.; qué relación guarda el autoconocimiento con el fenómeno de la autoadscripción mental y sus peculiaridades y cuál es el vínculo que une la cuestión epistémica con los conceptos de mérito y demérito. En segundo lugar, expondremos con detalle la postura de Bar-On: en qué consiste, qué novedades introduce su neoexpresivismo respecto de posiciones expresivistas clásicas, por qué da cuenta de la seguridad especial de nuestras autoadscripciones, etc., así como su explicación del autoconocimiento. En la cuarta sección, nos centraremos en presentar un argumento contra la explicación neoexpresivista del autoconocimiento que ofrece Bar-On, el cual hace uso del carácter meritorio que acompaña necesariamente a nuestras afirmaciones de conocimiento. En suma, sostendremos que nuestras autoadscripciones mentales no pueden considerarse afirmaciones meritorias y, por ende, tampoco merecen ser etiquetadas como *afirmaciones de autoconocimiento*.

2. EL VÍNCULO ENTRE EL AUTOCONOCIMIENTO Y EL MÉRITO EPISTÉMICO

Al hablar de autoconocimiento, es posible que se nos vengan a la mente preguntas del estilo: ¿nos conocemos a nosotros mismos?, ¿me conozco a mí mismo?, ¿sé cómo soy? Con estas preguntas (perfectamente legítimas) no logramos capturar con meticulosidad el fenómeno del que nos queremos hacer cargo aquí. En alguna medida puedo llegar a conocerme: si soy egoísta o generoso, engreído o humilde, un tórrido amante o, más bien, frío y apagado. La búsqueda de esta clase de autoconocimiento puede tener relevancia filosófica y sentido dentro de la vida del ser humano, pero no es esta clase de autoconocimiento la que nos interesa presentar aquí.

El problema del autoconocimiento, tal y como ha sido presentado tradicionalmente, es una cuestión acerca de si nuestras autoatribuciones o autoadscripciones de estados mentales presentes articulan algún tipo de conocimiento especial, privilegiado sobre nosotros mismos. Hablamos de estados mentales presentes en general, ya sean creencias u otras actitudes de actitud proposicional, estados de conciencia (dolores, hambre, calor...), etc.; frente a estados mentales pasados, porque el conocimiento de estos últimos radica en otros mecanismos epistémicos, como la memoria, testimonios de otras personas... , que no ponen de manifiesto una posición epistémica privilegiada.

La mayor parte de la discusión sobre el autoconocimiento gira en torno a la idea de que esta clase de autoconocimiento es de un tipo especial o privilegiado: se argumenta la existencia de una asimetría epistémica entre el autoconocimiento de nuestros estados mentales presentes y las demás clases de conocimiento (Boyle, 2009, 134 y ss., en especial, nota al pie 3; Moran, 2001, xxix; Rödl, 2007, vii y ss.; Wright, 1998, 101-102). Esta asimetría se pone de manifiesto una vez que nos percatamos de que rara vez estamos dispuestos a cuestionar la posición epistémica de los sujetos que se autoadscriben su estado mental; de que normalmente decimos de estos que poseen un conocimiento más seguro que el de los demás acerca de su propia condición mental. No ocurre lo mismo con nuestros rasgos de carácter o personalidad, así como con nuestras emociones: no siempre sabemos que somos unos egoístas, que nos gusta una persona, o que odiamos a nuestra madre; pues el conocimiento de nuestra vida psicológica muchas veces es objeto del autoengaño, del carácter inconsciente de ciertos estados mentales, o de razonamientos viciados por nuestros deseos o anhelos [*wishful thinking*].

El problema del autoconocimiento se centra, por tanto, en el *carácter epistémico* de ciertas preferencias sobre nuestra propia condición mental que en la tradición filosófica especializada han recibido el nombre de “*avowals*”. Por “*avowals*” debemos entender aquellas autoadscripciones psicológicas en primera persona del presente del indicativo que proferimos de forma espontánea,

esto es, sin pararnos a reflexionar sobre nuestro estado mental y a las cuales no llegamos aplicando un razonamiento inferencial —no inferimos sobre base alguna que nos encontramos en tal o cual estado mental—. Así, “Me duele (x)”, “Yo creo que...”, “Yo deseo que.../(x)”, “Tengo calor”, etc., son consideradas “*avowals*” o, como nosotros las llamaremos, autoadscripciones mentales en *perspectiva de primera persona*. Esta última puntualización es importante en la medida en que debemos distinguir las autoadscripciones en primera persona gramatical de aquellas otras realizadas desde la perspectiva de primera persona —las que hemos denominado “*avowals*”—. Nuestro objetivo aquí no es discutir en qué consiste la perspectiva de primera persona; no obstante, baste para nuestros propósitos tener en cuenta que la perspectiva de primera persona hace referencia a la relación especial e intransferible que un sujeto mantiene con sus propios estados mentales presentes, relación que a los demás nos es imposible mantener con sus propios estados mentales¹.

Ahora bien, cabe la posibilidad de que uno pueda autoadscribirse estados mentales presentes desde una perspectiva de tercera persona utilizando para ello la primera persona gramatical. Por ejemplo, tras una intensa sesión de psicoanálisis, mi psicoanalista podría concluir que estoy enamorado de mi madre, y yo, que confié plenamente en él, podría fiarme plenamente de su diagnóstico y decir de mí mismo que estoy enamorado de mi madre. Aunque utilice la primera persona gramatical, la perspectiva que asumo con mi deseo amoroso es una de tercera persona, es una relación que únicamente podría diferenciarse en grado, pero no en especie, de la que otros podrían mantener con dicho estado mental. Dicho con otras palabras, la relación con el amor que

¹Más tarde, nos ocuparemos precisamente de la explicación que la filósofa estadounidense Dorit Bar-On ha desarrollado para dar cuenta de esta relación especial e intransferible, la cual caracteriza la perspectiva de la primera persona y el supuesto privilegio que esta conlleva.

siento no presenta asimetría alguna con la relación que otros podrían mantener con él².

2.1. *El autoconocimiento y el fenómeno de la autoadscripción mental*

Con esta primera aproximación al fenómeno del autoconocimiento hemos atisbado la estrecha relación entre el autoconocimiento y la discusión de la perspectiva privilegiada de la primera persona: el conocimiento que articulan nuestras autoadscripciones mentales es uno que poseemos al ejercitar una relación privilegiada o especial con nuestros estados mentales. No es casualidad que la discusión del autoconocimiento se halle ligada a la consideración de la perspectiva de la primera persona. Parece una idea de sentido común que cada uno de nosotros se encuentra mejor situado, epistémicamente hablando, que los demás en lo que concierne a nuestros estados mentales presentes³.

Ahora bien, como ya hemos apuntado, atribuimos al sujeto la posesión de esta perspectiva o actitud privilegiada cuando realiza una atribución mental en primera persona, esto es, un *avowal*. La explicación de la perspectiva de primera persona está supeditada al fenómeno de la autoadscripción mental especial o privilegiada: la primera ha de articularse atendiendo a una correcta caracterización de dicho fenómeno. Abogar por una u otra relación de primera persona depende de cómo expliquemos las peculiaridades mostradas por nuestras autoadscripciones mentales. La función, la naturaleza, el carácter

²De aquí en adelante, hablaremos simplemente de “autoadscripciones mentales” para referirnos a aquellas realizadas desde la perspectiva de la primera persona, a no ser que se indique lo contrario.

³No obstante, merece la pena hacer notar que esta intuición no es hegemónica dentro de la reflexión filosófica. En su libro *The Concept of Mind* (1949/2009), G. Ryle sostiene que no hay una diferencia substantiva entre el conocimiento que tenemos de otros y el conocimiento que poseemos acerca de nosotros mismos; únicamente difieren en el garante epistémico, la base de evidencia, que el sujeto posee en cada caso, yo poseo más evidencia acerca de mí mismo que otros, puesto que, por ejemplo, tengo acceso a mis soliloquios mudos (cf. Ryle, 1949/2009, p. 150-151).

epistémico, etc., de los que nos valgamos para explicar este fenómeno pondrán cercos a nuestra explicación de la perspectiva de primera persona. Asimismo, una explicación adecuada del autoconocimiento no puede obviar las características de este fenómeno, en la medida en que el autoconocimiento es conocimiento especial o privilegiado, conocimiento que el sujeto posee cuando se encuentra en esta posición privilegiada y que se manifiesta a través de dichas autoadscripciones.

Crispin Wright (1998) ha expuesto las peculiaridades del fenómeno de la autoadscripción mental [*the phenomenon of avowal*], el fenómeno de la autoadscripción mental con autoridad y que no es resultado de un proceso inferencial. Wright ofrece una caracterización de estas peculiaridades, a saber, de la autoridad, la transparencia y la ausencia de base evidencial, para los diferentes tipos de autoadscripciones mentales: las autoadscripciones fenomenológicas y de actitud proposicional—.

De acuerdo con Wright, si alguien realiza una autoadscripción mental en perspectiva de primera persona, parece que está fuera de lugar preguntarle por la evidencia sobre la cual descansa dicha afirmación: el sujeto sería incapaz de decir algo con sentido, de ahí que Wright concluya que, “en ese sentido, no hay nada sobre lo cual tales afirmaciones se apoyen”(Wright, 1998, 102). En esto consiste precisamente la ausencia de base evidencial [*groundlessness*] de nuestras autoadscripciones mentales. Por otro lado, la *presunción de autoridad* de nuestras autoadscripciones tiene que ver con la idea de que estas son las más indicadas para corregir las afirmaciones que otros hagan sobre nuestra condición mental: la comprensión del lenguaje en el que uno realiza la autoadscripción sumado a la sinceridad con la que es proferida la autoadscripción bastan para garantizar la verdad de lo que digo. Para Wright, nuestras autoadscripciones fenomenológicas —aquellas que tratan sobre nuestros estados de conciencia— poseen una autoridad mucho más fuerte que aquellas otras que versan sobre nuestras actitudes proposicionales, ya que estas últimas pueden estar sujetas

a casos de pensamiento ilusorio o autoengaño⁴. Por último, parece muy improbable que un sujeto que se encuentre en determinado estado mental, no tenga constancia de este, es decir, se nos antojaría absurdo que el sujeto profiriese ignorancia, que no estuviese dispuesto a aseverar que se encuentra en dicho estado mental. Esto es lo que Wright entiende por transparencia [*transparency*]: “típicamente, hay algo de absurdo en una afirmación del tipo: ‘No sé si P’” (1998, 103), —donde “P” es una autoadscripción mental—.

La idea central de Wright es que una explicación del autoconocimiento en este tipo de casos está condicionada por la naturaleza especial de esta clase de autoadscripciones mentales (cf. Wright, 102). En resumen, según Wright, al abordar el fenómeno del autoconocimiento hemos de tener en cuenta la caracterización del privilegio de la primera persona, privilegio que debe plasmarse también en el *carácter epistémico* de las atribuciones, especificando claramente en qué consiste la asimetría epistémica de la que nuestras autoadscripciones mentales hacen gala: por qué son especialmente seguras, por qué no solemos estar equivocados al atribuirnos nuestra condición mental, etc.

Hasta ahora, nuestra discusión se ha centrado en poner en relación la perspectiva de la primera persona con el fenómeno de la autoadscripción mental y el autoconocimiento. Nuestra idea principal, siguiendo a Wright, es que una explicación del autoconocimiento no puede hacer caso omiso de las características mostradas por nuestras autoadscripciones mentales, las cuales proferimos al encontrarnos en una posición privilegiada respecto de nuestros estados mentales presentes. Además, hemos puesto el énfasis en la presencia de una *asimetría epistémica* entre las perspectivas de la primera y la tercera

⁴Para Wright, “Me duele” constituye un criterio más sólido que “Creo que *Juego de Tronos* es una gran serie” para corregir las afirmaciones de los demás sobre mi estado mental, dado que nuestra autoatribución de creencia puede ser, por ejemplo, un caso de autoengaño. En dicho caso, *Juego de Tronos* no sería una serie de mi agrado, únicamente me atribuiría dicha creencia debido a que temo inconscientemente decepcionar a todo el mundo que me rodea y alaba la susodicha serie.

persona, asimetría heredada de la posición privilegiada o especial en la que uno se encuentra cuando lleva a cabo autoadscripciones mentales. En la siguiente sección nos preocuparemos por traer a colación el tema central que articula nuestra discusión sobre el autoconocimiento, a saber: la noción de *mérito* (y *demérito*) epistémico. La importancia de la noción de mérito *epistémico* saldrá a relucir en cuanto abordemos la explicación que Bar-On ofrece de la relación especial que los sujetos guardan con sus estados mentales en primera persona y cómo Bar-On da cuenta del autoconocimiento.

2.2. *El autoconocimiento y el mérito epistémico*

En la sección anterior cerramos la discusión atendiendo a la relación entre el autoconocimiento y el fenómeno de la autoadscripción mental, haciendo hincapié en que una explicación del autoconocimiento tenía que respetar las características del fenómeno —i.e., presunción de autoridad, ausencia de fundamentos [*grounds*] y transparencia—. En lo que sigue nos centraremos más bien en el carácter meritorio de las autoadscripciones en cuestión. Nuestro análisis de las afirmaciones de conocimiento es independiente del análisis tradicional del conocimiento según el modelo *CVJ* [*JTB*]: conocimiento es creencia verdadera justificada; no buscamos continuar el análisis tripartito del conocimiento añadiendo o sustrayendo condiciones necesarias y suficientes del propio concepto de conocimiento.

Nuestra intuición aquí es la siguiente: el mérito epistémico regula nuestras atribuciones de conocimiento, esto es, forma parte de las normas que regulan los usos epistémicos de preferencias lingüísticas como “Sé que...”, “Él sabe que...”, etc⁵. Cuando atribuimos conocimiento a alguien o a nosotros mismos no nos interesa simplemente la verdad de la proposición atribuida: las atribuciones de conocimiento no siempre ponen el acento sobre la adecuación

⁵Nuestra estrategia busca hacer justicia al carácter normativo que regula nuestras atribuciones de conocimiento (cf. Chrisman, 2010; 2012, sec. 3-4)

entre una determinada proposición —hacia la cual el sujeto tendría una determinada actitud cognitiva que representa el mundo de cierta manera— y la realidad. Cuando atribuimos conocimiento, no tenemos por qué tratar meramente de describir una relación que se da entre el sujeto y una proposición, sino que hacemos algo más. Ese algo tiene que ver con una dimensión normativa, y consiste, en gran medida, en la atribución de mérito o demérito atendiendo a lo que el sujeto ha llevado a cabo para merecer una evaluación epistémica positiva o negativa. Hablar de mérito y demérito realza otro aspecto de la idea de conocimiento que consideramos central: el hecho de que el conocimiento es algo de lo que el sujeto es responsable. Atribuimos a los sujetos conocimiento o desconocimiento de que *p* teniendo en cuenta si lo merecen o no: conocer es algo de lo que los sujetos se han de hacer cargo, y ese hacerse cargo es precisamente lo que queremos remarcar con la idea de mérito epistémico.

En *Intention* (§8), al abordar el conocimiento que poseemos sobre la posición de nuestras extremidades y de nuestros estados mentales intencionales, G. E. M. Anscombe escribe lo siguiente:

Pero digo que la [la posición de nuestros miembros] *sabemos* y no simplemente que *podemos indicarla* porque existe la posibilidad de acertar o equivocarnos: resulta importante hablar de conocimiento únicamente cuando se presenta un contraste entre “él *sabe*” y “él (solamente) *cree* que sabe”. [...] Por lo tanto, llamo “conocimiento” a este tipo de capacidad de expresión y no *solamente* “capacidad de expresión”. (Anscombe, 1957/1991, 57) [énfasis de la autora; mi añadido].

En estas consideraciones epistemológicas, Anscombe parece apuntar precisamente a esta misma idea. Solo donde hay una diferencia entre errar y acertar puede tener sentido el discurso epistémico: lo importante aquí es que estar en lo cierto o equivocarse es algo por lo que el sujeto merece loa y reproche, respectivamente. De ahí que Anscombe concluya el párrafo

contrastando tener conocimiento con algo que simplemente estamos inclinados a decir.

En una línea parecida se pronunciaba el filósofo austríaco Ludwig Wittgenstein en las *Investigaciones filosóficas*: “Esto es correcto: tiene sentido decir de otros que están en duda sobre si yo tengo dolor; pero no decirlo de mí mismo” (1953/2010, §246) o “(Y ‘saber’ quiere decir aquí que la expresión de incertidumbre carece de sentido)” (1953/2010, §247). En la segunda parte de las *Investigaciones*, Wittgenstein también afirma:

“Sé...” puede querer decir “No dudo...” —pero no quiere decir que la frase “Dudo...” sea *carente de sentido*, que la duda esté excluida lógicamente.

Decimos “Sé” cuando podemos decir “Creo” o “Supongo”; cuando podemos convencernos de algo. (Pero quien me objete que a veces decimos “¡Sin duda debo saber si siento dolor!”, “Sólo tú puedes saber lo que sientes” y cosas parecidas, debe tomar en consideración las ocasiones y la finalidad de estos modos de hablar. [...]). (1953/2010, 505) [énfasis del autor]

O en el *Cuaderno azul*:

Es decir, el caso en que tenderíamos a contestar a la pregunta: “¿Está usted seguro de que es esto lo que desea?”, diciendo: “Sin duda, yo tengo que conocer lo que deseo”. Ahora bien, comparemos esta respuesta con la que la mayoría de nosotros daríamos a la pregunta: “¿Conoce usted el alfabeto?”. La afirmación categórica de que lo conocemos ¿tiene un sentido análogo al de la afirmación anterior? En cierto sentido ambas afirmaciones eliminan la cuestión. Pero la primera no quiere decir: “Claro que conozco una cosa tan sencilla como ésta”, sino más bien: “Lo que usted me ha preguntado no tiene sentido”. Podríamos decir: en este caso adoptamos un mal

método de eliminar la cuestión. “Naturalmente que lo sé” podría reemplazarse aquí por “Naturalmente, no hay duda” e interpretarse que esto significa: “En este caso no tiene sentido hablar de duda”. (1958/2007, 61)

Wittgenstein, a diferencia de Anscombe, introduce la cuestión haciendo uso de la certidumbre e incertidumbre, cómo estas quedan emplazadas en los intercambios lingüísticos donde nuestras afirmaciones de conocimiento tienen lugar. Normalmente, cuando decimos de alguien que sabe que p —o este dice de sí mismo que sabe que p —, ‘accionamos’ un mecanismo lingüístico con el que dejamos de lado la consideración de la duda o incertidumbre. Esto no quiere decir que la duda carezca de sentido o que no tenga sentido alguno dudar de lo que decimos: podríamos haber pasado por alto cierta evidencia relevante que, de ser descubierta, revocase indefectiblemente nuestro conocimiento y hubiese mostrado que no *conocíamos* que p , sino que simplemente *creíamos* conocer que p . Wittgenstein apunta a la idea de que el juego de lenguaje normal en el que ponemos en circulación nuestro vocabulario epistémico es aquel en que tiene sentido decir que *podemos convencernos de algo*.

Lo importante de nuestros usos epistémicos normales no es que sean el resultado de un proceso en el que llegamos a convencernos de algo. Nuestro uso correcto del vocabulario epistémico no se agota en aquellos casos donde nuestro juicio de que p descansa sobre nuestra consideración de ciertas razones a favor o en contra de p . En último término, no hemos de confundir las particularidades del juego epistémico, las normas que gobiernan nuestras atribuciones de conocimiento con el proceso o medio por el cual llegamos a formarnos cierta creencia o juicio.

Las afirmaciones de conocimiento meritorias son aquellas en las que la duda queda eliminada —aunque no lógicamente—, aquellas que son el resultado de una *acción adecuada*, no fortuita y azarosa, del sujeto por eliminar todo atisbo de incertidumbre. El demérito viene dado por la incapacidad del sujeto

para dejar de lado la incertidumbre, la duda, en los intercambios lingüísticos habituales donde llevamos a cabo atribuciones de conocimiento. Estos casos se diferencian radicalmente de aquellos otros donde no tiene sentido dudar, donde la duda queda excluida lógicamente. Poder llegar a convencerse es el criterio que determina la puesta en escena de la evaluación epistémica; el mérito o el demérito tienen que ver con lo que el sujeto haga por excluir la incertidumbre de la que nuestras afirmaciones de conocimiento *pueden* ser objeto.

En la sección anterior presentamos la discusión acerca del estatus epistemológico de nuestras autoadscripciones mentales. Si aceptamos que dichas autoadscripciones pueden considerarse con todo derecho afirmaciones de autoconocimiento y no simples expresiones del estado mental del sujeto —como veremos más adelante al caracterizar la propuesta de Bar-On—, hemos de situarlas dentro del marco epistémico (normativo) en el que se encuentran nuestros usos cotidianos, normales, de afirmaciones de conocimiento. El reto, por tanto, consistirá en evaluar el expresivismo de Bar-On respecto de las autoadscripciones atendiendo a esta nueva dimensión de evaluación, a saber: la del mérito (y demérito).

Nuestro objetivo principal es ofrecer una evaluación del fenómeno del autoconocimiento tal y como ha sido presentado por Dorit Bar-On desde su postura neoexpresivista. Por ello, dedicaremos la siguiente sección a presentar la visión expresivista de Bar-On, las coordenadas que ocupa dentro del plano expresivista y la estrategia que Bar-On pone en juego para ligar su expresivismo con la clase de autoconocimiento discutida anteriormente.

3. EL NEOEXPRESIVISMO DE DORIT BAR-ON

Dorit Bar-On ha presentado a lo largo de estos años (2004, 2009, 2011, 2015...) una versión *híbrida*⁶ de expresivismo que intenta separarse de las

⁶El carácter híbrido del expresivismo de Bar-On se debe a que conjuga dos nociones de expresión bien diferenciadas que permiten hacer hueco a una semántica

posiciones expresivistas y emotivistas que buscan dar cuenta de la significación [*meaningfulness*] de nuestras autoadscripciones mentales y éticas acudiendo, respectivamente, al estado mental o las actitudes a favor o en contra que dichas autoadscripciones expresan —negando, además, la posibilidad de que estas puedan ser evaluadas conforme a la verdad—. Bar-On se aparta de aquellas propuestas no-cognitivistas que se proponen explicar la semántica de las susodichas autoadscripciones sin postular proposiciones que representen ciertos estados de cosas. En general, su propuesta se configura como una visión no-ideacionista, no-psicologista, de la semántica de las autoadscripciones mentales, ya que no pretende reducir la explicación del significado de ciertos ítems del lenguaje, viz. autoadscripciones mentales, enunciados éticos, etc., a la mención de determinados estados mentales. (v. Bar-On y Sias, 2013, 705, 709).

Aquí examinaremos principalmente la aplicación de esta estrategia semántica al discurso mental. En lo que concierne al discurso mental, la propuesta de Bar-On constituye una actualización de las explicaciones expresivistas clásicas o tradicionales de la especial autoridad de nuestras autoadscripciones mentales, según las cuales la autoridad es explicada por el carácter expresivo y no-aseverativo, de dichas autoadscripciones. Bar-On no está dispuesta a aceptar, al igual que el expresivismo clásico, que nuestras autoadscripciones no puedan ser evaluadas respecto a su valor de verdad —esto

composicional y explicar, al mismo tiempo, las peculiaridades de nuestros actos de habla. Las nociones de s-expresión (expresión en el sentido semántico) y a-expresión (expresión en el sentido de acción) (v. *infra*) permiten, con independencia de dominio lingüístico en que nos situemos, dar cabida a la idea de que nuestras preferencias lingüísticas, como productos, expresan proposiciones veritativo-evaluables y, como actos, expresan estados mentales, actitudes a favor de o en contra de *x*, etc.

No obstante, a diferencia del neoexpresivismo de Bar-On, el hibridismo de otras posturas expresivistas se caracteriza más bien por la naturaleza de los estados mentales expresados a través de las diferentes expresiones normativas: frente a posturas expresivistas *puras*, las cuales consideran que la naturaleza de los estados mentales expresados es o bien cognitiva, o bien no-cognitiva; versiones híbridas del expresivismo han defendido que la condición mental expresada tiene componentes cognitivos y no-cognitivos (v. Bar-On y Sias, 2013, 701 y ss.)

es, negar que posean contenido veritativo-evaluable—. La razón principal de este rechazo es que dicha postura hace de la continuidad semántica entre autoadscripciones mentales como “Me duele” y otras expresiones lingüísticas como, por ejemplo, “Él tiene dolor”, “Me dolía”, “Si me duele, iré al médico”, etc., una cuestión problemática, ya que la existencia de contenido veritativo-evaluable queda comprometida.

Bar-On sostiene que el expresivismo clásico acertó al indicar el carácter expresivo de ciertas preferencias mentales, pero que ello no conlleva postular una discontinuidad semántica con otras oraciones significativas. Para ella, el expresivismo ha logrado captar la relación entre nuestros estados mentales y ciertas expresiones lingüísticas y no-lingüísticas, pero no ha tenido en cuenta la naturaleza de ciertos vehículos expresivos, que marca la diferencia entre la expresión lingüística y autoadscriptiva propia de los humanos entrenados en una lengua y la expresión de estados mentales llevada a cabo por bebés y otros animales no-humanos que carecen de lenguaje con propiedades semánticas (cf. Bar-On, 2004, 241 y ss.; 2015, 140-141 y nota 27). Esta idea es la que intenta plasmar Bar-On diferenciando entre dos sentidos de expresión: *a-expresión* —expresión en sentido de acción— y *s-expresión* —expresión en sentido semántico— (v. Bar-On, 2004, 248; 2009, 67 y ss.; 2011, 194; 2015, 140). Cuando un sujeto *a-expresa*, este está manteniendo una relación con su estado mental actual y un vehículo expresivo, de tal manera que lo que hace es dar voz, exhibir, mostrar, su propio estado mental a través de cierto vehículo, esto es, *a-expresar* consiste en dar voz al estado mental haciendo algo intencionalmente; por otro lado, *s-expresar* consiste en una relación semántica entre el vehículo de expresión y su contenido semántico (cf. Bar-On, 2015, 140). Para Bar-On, son los productos de nuestras preferencias lingüísticas los que poseen propiedades semánticas, los que *s-expresan* una proposición que constituye el significado de la expresión lingüística proferida.

Esta distinción permite a Bar-On salvar la continuidad semántica de nuestras autoadscripciones mentales, dado que estas *s-expresarían* (al igual

que el resto de oraciones en un lenguaje natural) proposiciones con sus correspondientes condiciones de verdad; y, al mismo tiempo, capturar la idea de que nuestras preferencias lingüísticas, como actos, a-expresan el estado mental en que nos encontramos. El estado mental en que se encuentra el sujeto que a-expresa tiene que ver con las normas que regulan nuestras preferencias mentales en primera persona, entendidas como actos, y no con la determinación del contenido semántico de los productos lingüísticos generados. En los últimos desarrollos de su postura, Bar-On ha expresado esta idea afirmando que encontrarnos en el estado mental que a-expresamos constituye una condición para llevar a cabo *con éxito* preferencias de autoadscripciones mentales en perspectiva de primera persona (v. Bar-On y Chrisman, 2009, 146; Bar-On y Sias, 2013, 707).

Al escindir del contenido semántico el estado mental a-expresado, el neoexpresivismo de Bar-On recupera la verdad y la falsedad como etiquetas evaluativas legítimas. Bar-On no solo permite que la proposición s-expresada sea falsa, sino que entiende que el sujeto puede ser víctima de fallos expresivos (v. Bar-On, 2015, 145-6). El sujeto puede fallar a la hora de a-expresar su estado mental cuando no se encuentra en el estado mental que se autoadscribe, violando así las condiciones de éxito para la realización de autoadscripciones mentales, que, como consecuencia, da lugar a autoadscripciones mentales ‘de mentira’, solo en apariencia (v. Bar-On y Chrisman, 2009, 146).

En suma, el neoexpresivismo de Bar-On suscribe la tesis expresivista —i.e., nuestras preferencias mentales expresan nuestro estado mental—, pero no la postura semántica, no-cognitivista, que el expresivismo clásico enarboló. Bar-On piensa que existen grandes similitudes entre nuestras expresiones lingüísticas y otro tipo de conductas expresivas no-lingüísticas: ambas a-expresan la condición mental del ser que lleva a cabo cierta acción. No obstante, el neoexpresivismo de Bar-On rechaza que debamos admitir que nuestras autoadscripciones mentales —ítems lingüísticos bien formados, aprendidos y significativos— expresan el estado mental de un sujeto en el

mismo sentido en que lo hace una expresión natural, no-lingüística (v. Bar-On, 2004, 286). Un sujeto que haya dominado una determinada lengua, que haya aprendido a hacer autoadscripciones mentales, habrá adquirido un modo especial de expresar sus estados mentales: un modo de expresión autoadscriptivo (v. Bar-On, 2004, 317; 2015, 142), de ahí que Bar-On haya recalcado que la capacidad expresiva de los humanos entrenados en una lengua no ha de ser considerada meramente un logro psicológico, sino también lingüístico (cf. 2015, 145), dado que nuestra capacidad para expresar autoadscriptivamente nuestros estados mentales no pone en juego simplemente nuestra capacidad para hacerlos manifiestos, sino también nuestra capacidad para expresar proposiciones acerca de nuestra propia vida mental.

3.1. *Bar-On: neoexpresivismo y autoconocimiento*

Más arriba comentábamos cuál era la tarea fundamental de una explicación del autoconocimiento de acuerdo con C. Wright (1998). Según este, una explicación del problema del autoconocimiento debería dar cuenta de las características que presenta el fenómeno de la autoadscripción mental en perspectiva de primera persona, explicando cuál es la posición especial en la que se encuentra el sujeto que realiza tales autoadscripciones (v. 1998, 102). Un proyecto que persigue esta estrategia puede recurrir a nociones epistémicas —como la de sentido interno, instrospección, etc.— para explicar por qué el sujeto se encuentra en una posición de privilegio al proferir las autoadscripciones mentales, al mismo tiempo que otorga a dichas autoadscripciones la etiqueta que las certifica como afirmaciones de autoconocimiento privilegiado. Una estrategia de este estilo busca poner de manifiesto las particularidades epistémicas que hacen de las autoadscripciones mentales en perspectiva primera persona afirmaciones de autoconocimiento, al mismo tiempo que explica en qué consiste el privilegio o autoridad de la primera persona.

En cambio, Bar-On se muestra reacia a asumir una estrategia de este estilo (v. 2015, 135). Para Bar-On, explicar por qué nuestras autoadscripciones mentales son afirmaciones de autoconocimiento privilegiado no es el *datum* por el que alguien interesado en dar cuenta del fenómeno de la autoadscripción mental deba preocuparse en primera instancia. La explicación de las características de nuestras autoadscripciones mentales debería comenzar sin hacer uso de nociones que exploten su supuesta dimensión epistémica. Tomar como *datum* principal la explicación del autoconocimiento, sostiene Bar-On, presenta problemas derivados de una mala comprensión de la especial seguridad que nuestras autoadscripciones mentales poseen (v. e.g. Bar-On, 2009; 2015). La especial seguridad, sostiene Bar-On, no tiene que ver con la dimensión epistémica de nuestras autoadscripciones mentales, sino con su carácter expresivo.

Bar-On, por tanto, entiende que una explicación del privilegio que un sujeto manifiesta al autoadscribirse su condición psicológica no ha de supeditarse a la explicación del autoconocimiento. Por esto mismo, la estrategia de Bar-On (cf. 2004, 341-342) es bipartita: en primer lugar, busca elaborar una explicación de por qué nuestras autoadscripciones mentales están provistas de una seguridad especial —una explicación que no haga uso de ningún método o capacidad epistémica gracias a la cual el sujeto sepa cuál es su condición mental— y, de este modo, explicar la autoridad que caracteriza a este tipo de adscripciones. En segundo lugar, tras rechazar una solución deflacionista al problema del autoconocimiento —aunque podría endorsarla (cf. Bar-On, 2004, 345)—, según la cual nuestras autoadscripciones mentales carecerían de una dimensión epistémica y, por ende, no podrían ser consideradas afirmaciones de conocimiento, Bar-On se propone dar una explicación adecuada de cómo esta clase de autoadscripciones pueden ejemplificar casos de autoconocimiento privilegiado.

En lo que respecta al reto de la seguridad especial, Bar-On sostiene que la fuerte presunción de verdad de nuestras autoadscripciones mentales,

su resistencia a los retos epistémicos, la confianza que otros muestran hacia ellas, esto es, que raramente otros cuestionen la veracidad de las susodichas autoadscripciones —en resumen, lo que hemos denominado la autoridad o el privilegio de la primera persona—, no descansa sobre creencias de segundo orden que el sujeto se haya formado haciendo uso de algún método epistémico especial e intransferible. La seguridad especial de nuestras proferencias mentales se debe, según Bar-On, a que estas a-expresan —dan voz, muestran, exhiben— *los mismos estados mentales que nuestras autoadscripciones nos atribuyen*: los mismos estados mentales con los que la proposición s-expresada por la oración proferida nos relaciona. Esto explica que normalmente dichas autoadscripciones sean consideradas verdaderas, que nadie intente corregirnos o preguntarnos cómo es que sabemos que nos encontramos en el estado mental autoadscrito. El carácter a-expresivo de nuestras autoadscripciones mentales también explica la ausencia de fundamentos o razones [*groundlessness*] para autoadscribirnos estados mentales, dado que nuestra capacidad expresiva no pone en juego ninguna facultad epistémica a través de la cual justificar u obtener razones a favor de nuestras autoadscripciones mentales. Del mismo modo, en tanto que nos autoadscribimos el mismo estado mental que a-expresamos, es decir, dado que *decimos* en qué estado mental nos encontramos, carecería de sentido cualquier expresión de duda sobre nuestra propia condición mental [*transparency*].

Pero, ¿qué nos permite hablar de autoconocimiento? Bar-On introduce el asunto considerando las razones que hablan a favor de apoyar una postura deflacionista sobre el autoconocimiento. De acuerdo con la filósofa estadounidense, el deflacionismo acerca del autoconocimiento se nutre de la idea de que el autoconocimiento ha de tener una fuente epistémica distinta (cf. Bar-On, 2004, 344) —e.g., la introspección, o algún tipo de sentido interno—. Bar-On atribuye a dicho deflacionismo la presuposición de que el autoconocimiento genuino de la primera persona viene dado por un método epistémico adecuado para obtener información sobre los hechos relevantes

—i.e., nuestros estados mentales presentes—. En última instancia, Bar-On sostiene que el deflacionismo sobre el autoconocimiento se nutre de una concepción recognoscitiva [*recognitional*] del autoconocimiento.

El neoexpresivismo de Bar-On rechaza la idea de que el sujeto deba llevar a cabo un esfuerzo epistémico recognoscitivo —e.g., observar, hacer una inferencia, un juicio, etc.— con el fin de obtener autoconocimiento. Su neoexpresivismo niega que el autoconocimiento deba ser considerado un logro cognitivo o epistémico en ese sentido (v. Bar-On, 2004, 353). No obstante, en tanto que a-expresar nuestros estados mentales en el sentido que nos ocupa conlleva emitir autoadscripciones lingüísticas bien formadas con propiedades semánticas, dichas autoadscripciones requieren del dominio gramatical, lógico y epistemológico necesario para aplicar los conceptos pertinentes a la hora de realizar una a-expresión mediante vehículos lingüísticos. En este sentido, a-expresar a través de vehículos lingüísticos autoadscriptivos sí que conlleva un esfuerzo y logro cognitivo.

Llegados a este punto, cabría preguntarnos cómo se relacionan la especial seguridad de las autoadscripciones mentales, la ausencia de esfuerzo recognoscitivo y la posibilidad de conocimiento. En *Speaking my mind* (2004), Bar-On encaja las diferentes piezas de este rompecabezas retomando el asunto de la seguridad especial a la luz del carácter no-recognoscitivo de nuestras autoadscripciones. El carácter a-expresivo de nuestras autoadscripciones mentales indica precisamente la ausencia de una vía epistémica recognoscitiva a través de la cual averigüemos o descubramos nuestra condición mental. La estrategia de Bar-On consiste, pues, en hacer uso de un modelo de conocimiento no-recognoscitivo para las autoadscripciones mentales. Bar-On sostiene que las autoadscripciones mentales son inmunes a errores de identificación [*immune to error through misidentification*] y, además, inmunes a errores de adscripción [*immune to error through misascription*] (v. Bar-On, 2009, 64 y ss.; 2011, 191 y ss.; 2015, 143). La inmunidad de nuestras autoadscripciones mentales a tales tipos de errores se debe a que el sujeto no tiene razones para afirmar

que *algún otro* se encuentra en *cierto estado mental M'* independientemente de las que posee para afirmar o juzgar que es *él mismo* quien se encuentra en *el estado mental M* (Bar-On, 2011, 191; 2015, 143). En concreto, las autoadscripciones mentales, sostiene Bar-On, no son inmunes a todo tipo de errores, sino a aquellos errores de identificación y adscripción relativos a nuestras capacidades reconocitivas (v. Bar-On, 2011, 193), dado que llevar a cabo una autoadscripción mental no pone en juego ningún mecanismo de este tipo.

Bar-On sugiere que deberíamos preguntarnos cómo el sujeto es capaz de obtener la información de que está en el estado mental M (con cierto contenido *c*, en caso de poseerlo). Bar-On ofrece una solución que no hace descansar la autoadscripción mental sobre mecanismos reconocitivos, ya sea para reconocer al 'portador' de los estados, ya para averiguar cuál es el estado mental en que se encuentra (y su contenido, en caso de tenerlo). La respuesta articulada por Bar-On acude a una *capacidad práctica* del sujeto que no se apoya en la consideración de las fuentes epistémicas propias de un juicio o creencia reconocitiva, una capacidad que cada sujeto solo puede ejercitar respecto de sus propios estados mentales: la capacidad para a-expresarlos de un modo autoadscriptivo, esto es, utilizando un vehículo lingüístico que s-exprese una proposición sobre él mismo y su condición mental; capacidad que, como ya hemos visto, garantiza la seguridad especial que nuestras autoadscripciones verbales presentan intuitivamente.

Este relato neoexpresivista le sirve a Bar-On para hacer hueco a una explicación del autoconocimiento que deje a un lado las presuposiciones deflacionistas y reconocitivas: una explicación que de cuenta del carácter epistémico de las autoadscripciones mentales respetando su carácter expresivo. Bar-On asume esta tarea recogiendo la influencia del famoso modelo *CVJ [JTB]*, i.e. conocimiento es creencia verdadera justificada. El requisito de la verdad es satisfecho en la medida en que nuestras autoadscripciones poseen una fuerte presunción de verdad, puesto que cuando nos autoadscribimos cierto estado

mental al a-expresarlo, *casi siempre* nos encontramos en ese estado mental, exceptuando casos extraños donde el sujeto falla al a-expresar su estado mental, violando las condiciones de éxito de las que hablábamos en la sección anterior.

Asimismo, Bar-On sostiene que, aunque el sujeto se encuentre a-expresando el estado mental en que se encuentra, puede decirse de este que posee una creencia de segundo orden sobre su propio estado mental, una creencia con el mismo contenido que la proposición s-expresada. Para Bar-On esta creencia no puede ser caracterizada en términos de una creencia opinativa [*opining sense*] (v. Bar-On, 2004, 363), puesto que cuando el sujeto se autoadscribe su condición mental no lo realiza sobre la base de ningún juicio que identifique o reconozca la evidencia relevante. Ahora bien, Bar-On afirma que existe un sentido de creencia que puede acomodar el ‘contacto cognitivo’ que las autoadscripciones representan. Bar-On denomina a este nuevo sentido de creencia “*holding-true sense*”, de acuerdo con el cual una persona poseería una creencia de que *p* si aceptase *p* al considerar *p*⁷.

No obstante, Bar-On defiende que podríamos atribuir al sujeto una creencia de segundo orden en otro sentido aún más fuerte. Bar-On denomina a este nuevo sentido de creencia “sentido autoadscriptivo” [*self-ascriptive sense*]. Según la filósofa estadounidense, podríamos atribuir una creencia de segundo orden a un sujeto simplemente por el hecho de que este lleve a cabo la autoadscripción mental relevante. El hecho de que el sujeto profiera intencionalmente una autoadscripción mental con cierto contenido semántico es suficiente para atribuir al sujeto una creencia de segundo orden. En resumen, Bar-On sostiene que nuestras autoadscripciones pueden representar una actitud cognitiva del sujeto —una a la que este no llega a través de cierta base reconocitiva— en la medida que nuestra capacidad para a-expresar

⁷Chrisman ha señalado lo absurdo que sería atribuir una creencia en el sentido de ‘tenerla por verdadera’ a alguien que se autoadscribe su estado mental (v. Chrisman, 2009, 7), puesto que en el caso de las autoadscripciones mentales el estado mental atribuido es uno en el cual el sujeto se encuentra, por lo que no parece convincente afirmar que es algo que el sujeto no ha considerado o teniendo en cuenta.

autoadscriptivamente los estados mentales en que nos encontramos involucra la producción de ítems lingüísticos con propiedades semánticas, esto es, ítems que transmiten información acerca del estado psicológico del hablante.

Aun así, el neoexpresivismo aún nos ha de explicar, según los criterios del modelo *CVJ* del que Bar-On se vale en su argumentación, qué convierte a la creencia verdadera que un sujeto tiene al a-expresar autoadscriptivamente el estado mental en que se encuentra en un caso genuino de autoconocimiento privilegiado. Dicho con otras palabras, tendría que ofrecer, según el modelo *CVJ*, un relato consistente que permita hablar de justificación para las creencias verdaderas que un sujeto posee sobre su estado mental a-expresado.

El rechazo radical de la presuposición deflacionista sobre el autoconocimiento (v. supra) lleva a Bar-On a ofrecer varias posturas *compatibles* (Bar-On, 2011, 190; 2015, 143) con su explicación neoexpresivista de la especial seguridad de las autoadcripciones mentales (cf. 2004, 369-396). La explicación a la que Bar-On da más peso entre todas las expuestas en su libro hace uso de un concepto de justificación disyuntivista. Bar-On entiende que el estado mental que hace verdadera a la proposición s-expresada también justifica la creencia de segundo orden que posee el sujeto. Según Bar-On, la razón por la que el sujeto lleva a cabo un acto de expresión (la causa racional), a saber, su propio estado mental, es también aquello que justifica la autoadcripción mental entendida como producto. En este sentido, puede decirse que el sujeto realiza un logro epistémico, pese a que, como ella misma admite, esto no requiera un esfuerzo epistémico por parte del sujeto (Bar-On, 2004, 390), puesto que el esfuerzo es más bien *expresivo*.

Esta maniobra permite a Bar-On defender que las autoadcripciones mentales en primera persona representan casos de autoconocimiento genuino y, además, especial o único, ya que se trata de una clase conocimiento que otro no puede poseer. En la medida en que aquello que justifica la creencia del sujeto no es otra cosa que su propio estado mental, se trata de un conocimiento que únicamente este puede poseer. Su carácter privilegiado reside en que en los actos

expresivos el estado mental en que el sujeto se encuentra es al mismo tiempo la razón para su acción y la razón epistémica que justifica su creencia de segundo orden (v. Bar-On, 2004, 391).

4. NORMATIVIDAD EPISTÉMICA Y AUTOADSCRIPCIONES MENTALES

Esta última sección la dedicaremos a conectar nuestras reflexiones normativas iniciales con la cuestión del autoconocimiento tal y como es presentada por Bar-On con el fin de respaldar la hipótesis central de este trabajo, a saber: que las autoadscripciones mentales no poseen dimensión epistémica, dado que carece de sentido evaluarlas atendiendo al mérito o demérito. Nuestra argumentación descansaba principalmente sobre una distinción entre los mecanismos epistémicos gracias a los cuales obtenemos conocimiento y el espacio normativo dentro del que se inscriben nuestras afirmaciones de conocimiento. En este sentido, mientras la explicación de Bar-On aborda principalmente las cuestiones de la justificación y la fuente epistémica de nuestras autoadscripciones mentales, nosotros estudiaremos más bien por qué estas *merecen* la etiqueta de “autoconocimiento”.

Nuestra tesis principal es que aquello que hace de nuestras preferencias lingüísticas afirmaciones de conocimiento es su carácter meritorio, el cual está íntimamente conectado con lo que el sujeto hace al proferir una afirmación de conocimiento, a saber, eliminar la cuestión de la duda. Asimismo, al hablar sobre dejar la duda a un lado, aclaramos que esto podría hacerse siempre y cuando tuviese sentido que el sujeto pudiese llegar a convencerse sobre algo, lo cual no equivalía a afirmar que el sujeto había llegado a conocer que *p* considerando las razones a favor de *p*. En resumen, las afirmaciones de conocimiento tenían sentido allí donde la consideración de las razones para afirmar cierta proposición fuese pertinente. Esto no equivalía a sostener que únicamente podíamos atribuir conocimiento a un sujeto cuando este se formase

una creencia atendiendo a la evidencia o razones disponibles para adoptar tal compromiso cognitivo.

Por esto mismo, no discutiremos el concepto de justificación que maneja Bar-On, o sus opiniones acerca de en qué sentido el sujeto posee una creencia, sino que analizaremos si las autoadscripciones mentales, las cuales realizamos desde una perspectiva de primera persona con una autoridad especial, sin fundamentos o razones y con transparencia, merecen ser denominadas “afirmaciones de conocimiento”. Para ello, estudiaremos las autoadscripciones mentales bajo el prisma de los requisitos normativos que gobiernan nuestros intercambios lingüísticos, en concreto, uno de los requisitos que hemos considerado fundamental en el marco de las afirmaciones de conocimiento, viz. el *carácter meritorio* de las afirmaciones de conocimiento. En consecuencia, presentaré un argumento a favor de la tesis de que nuestras autoadscripciones mentales no son afirmaciones de conocimiento, y lo discutiré a la luz de diferentes objeciones y dudas que el mismo pueda sugerir. El argumento es el siguiente:

P1 Respecto de toda afirmación de conocimiento tiene sentido preguntarse por las razones que uno posee para aseverarla.

P2 No tiene sentido preguntarse por las razones de nuestras autoadscripciones mentales en primera persona.

C Nuestras autoadscripciones mentales no son afirmaciones de conocimiento.

Este argumento viene motivado por un fragmento de *El cuaderno azul*, donde Wittgenstein se pregunta por la diferencia entre los casos de autoconocimiento y los de conocimiento de otras mentes. Reza así:

Compárense los dos casos: 1. “¿Cómo sabe usted que *él* tiene dolores?” “Porque le oigo quejarse”. 2. “¿Cómo sabe usted

que usted tiene dolores?” “Porque los *siento*”. Pero “yo los siento” significa lo mismo que “yo los tengo”. Por tanto, esto no ha constituido ninguna explicación. (Wittgenstein, 1958/2007, pág. 105) [énfasis del autor]

Wittgenstein pone de manifiesto que en los hipotéticos casos de autoconocimiento, a diferencia de aquellos casos donde conocemos la mente de los demás, el sujeto no dispone de razones genuinas para apoyar su autoadscripción mental. Aunque en este fragmento Wittgenstein se centre en autoadscripciones de estados de conciencia o fenomenológicos, este argumento se podría extrapolar sin problemas a las autoadscripciones de actitud proposicional. Como Wright indica, las autoadscripciones mentales genuinas de actitud proposicional no son el resultado de un proceso de autointerpretación, no descansan sobre procesos autointerpretativos, sino que más bien constituyen la evidencia para llevar a cabo estos procesos (Wright, 1998, 104). En este sentido, cuando llevamos a cabo una autoadscripción de actitud proposicional genuina no poseemos evidencia o razones a favor de ella.

Con nuestro argumento nos proponemos poner en relación el carácter meritorio que hemos atribuido a las afirmaciones de conocimiento a través de su papel para ‘disolver’ la duda y la cuestión del autoconocimiento. Recordemos que este papel solo era llevado a cabo por nuestras afirmaciones de conocimiento allí donde tuviese sentido decir que puede existir un convencimiento —en este caso del propio sujeto— o, lo que es lo mismo, que el sujeto *podría* participar de un proceso por medio del cual aportar razones a favor de su afirmación. Ahora bien, el objetivo de nuestro argumento anterior no es otro que denunciar el estatus epistemológico de nuestras autoadscripciones mentales: no tiene sentido considerarlas afirmaciones de autoconocimiento, dado que no pueden ser enmarcadas correctamente dentro del espacio epistémico, no tiene sentido hablar de duda o de convencimiento cuando nos autoadscribimos un estado mental en perspectiva de primera persona. Llevar a cabo una autoadscripción

mental no es un método que el sujeto utilice para eliminar una hipotética incertidumbre acerca de su condición mental, la autoadscripción no es un método para eliminar la duda, ni tiene sentido hablar de razones a favor de dicha autoadscripción.

A diferencia de nosotros, Bar-On hace de las cuestiones acerca del método o la facultad epistémica que el sujeto utiliza para conocer su condición psicológica y de la noción de justificación la piedra angular del debate epistemológico (cf. 2004, 367 y ss.). Bar-On denuncia una noción de justificación demasiado restrictiva, centrada en aquellos casos en los que el sujeto realiza una afirmación respaldada por una actitud cognitiva alcanzada a través de un proceso o esfuerzo cognitivo —e.g., una inferencia, un juicio basado en la evidencia, etc.—. Bar-On ofrece ejemplos de casos de conocimiento que no caen bajo esta etiqueta de justificación, en especial, el conocimiento perceptivo. Según Bar-On, cuando vemos un vaso roto no nos formamos la creencia perceptiva de que el vaso está roto sobre la base de ciertas razones, esto es, no llegamos a descubrir, averiguar, que el vaso está roto, sino que más bien captamos o aprehendemos tal hecho. ¿Por qué hay que suponer entonces —se pregunta Bar-On— que el conocimiento de mis estados mentales presentes es tal que es el resultado de un proceso de descubrimiento o averiguación?

Al presentar la postura de Bar-On, insistimos en que esta rechazaba la idea de que debamos poner en marcha un mecanismo epistémico con el fin de averiguar o descubrir nuestra condición mental y gracias al cual lleguemos a poseer creencias especialmente seguras. No obstante, nuestra crítica tiene que ver no tanto con el proceso en sí como con la función que cumplen nuestras afirmaciones de conocimiento y las reglas que las gobiernan. La analogía con los casos de percepción es equívoca por el siguiente motivo: que no lleguemos a conocer un hecho del mundo a través de ciertos mecanismos reconocitivos no quiere decir que la discusión sobre las razones para afirmar que el vaso está roto carezca de sentido. El carácter meritorio no descansa sobre los

mecanismos epistémicos que el sujeto pone en marcha, sino que forma parte de la normatividad del discurso epistémico. “Sé que el vaso está roto” puede llegar a ser una afirmación de conocimiento, aunque hayamos llegado a saber tal cosa teniendo acceso directo al hecho de que el vaso está roto. Alguien podría preguntarnos por qué sabemos que el vaso está roto y, aunque no sea algo que hayamos descubierto, estaríamos en posición para ofrecer diferentes razones, e.g., porque vemos cascotes de cristal a su alrededor, se han formado grietas por toda su superficie, etc. Esto no tiene sentido en el caso de las autoadscripciones mentales, tal y como veíamos en la anterior cita de Wittgenstein.

La especial seguridad de las autoadscripciones mentales, su autoridad, puede explicar por qué está fuera de lugar *preguntar* por las razones. Sin embargo, si queremos evaluarlas como afirmaciones de conocimiento, tenemos que preguntarnos si tiene sentido el recurso a las razones. Estamos planteando si las autoadscripciones deberían ser consideradas afirmaciones de autoconocimiento de acuerdo a ciertas normas que rigen la dimensión epistémica de nuestros intercambios dialécticos. Que un movimiento no esté legitimado, debido en este caso a la dimensión expresiva, no imposibilita que nos preguntemos por el estatus epistémico de dicha afirmación. Considerar nuestras autoadscripciones mentales como afirmaciones de autoconocimiento pasa por considerar su carácter meritorio, algo que Bar-On no consigue poner de manifiesto hablando de verdad, creencia y justificación en el caso de las autoadscripciones mentales.

Bar-On podría argumentar que estamos pasando por alto la posición epistémica del sujeto. En la síntesis llevada a cabo por Bar-On (v. Bar-On, 2004, 388 y ss.), donde busca ofrecer una explicación convincente del fenómeno del autoconocimiento aunando diferentes componentes presentados a lo largo del capítulo, esta retoma un elemento presentado en su ‘ruta intermedia hacia el autoconocimiento’ [*A Middle Road to Self-knowledge*] (Bar-On, 2004, 381-388). Según Bar-On, el sujeto que a-expresa su estado mental es inmune a los errores de identificación y adscripción en tanto que su autoadscripción

no descansa sobre mecanismos reconocitivos. Esto —dice la autora— deja al sujeto que a-expresa en una situación de “inocencia epistémica, y, por tanto, no se le considera responsable de las alternativas que harían falsa a su autoadscripción” (Bar-On, 2004, 390) [traducción propia]. Empero, que a-expresar deje al sujeto en esa posición indica precisamente que su ‘innocencia epistémica’, esto es, que no sea *responsable* de las cuestiones epistémicas de su autoadscripción, viola las reglas que gobiernan nuestras evaluaciones epistémicas. Que no sea responsable de las alternativas indica la ausencia del carácter meritorio necesario a la hora de atribuir conocimiento.

Como indica Wright, la inmunidad al error de identificación no tiene que ver con la naturaleza de los enunciados que poseen tal inmunidad, sino que más bien “depende del tipo de fundamento que el hablante tiene para el enunciado” (Wright, 2001, 327) [traducción propia]. Que un enunciado sea inmune al error de identificación quiere decir que en caso de que las razones que el sujeto posee para afirmar que el enunciado versa sobre él mismo fuesen echadas por tierra, entonces el sujeto no tendría razones para llevar a cabo siquiera una generalización existencial del tipo “Alguien tal-y-tal”. Esta idea puede extrapolarse fácilmente a la inmunidad al error de adscripción que propone Bar-On.

Ahora bien, nuestra intuición es la siguiente: la incapacidad para mantener en pie una generalización existencial ya sea acerca del sujeto o de su estado (y el contenido de este, en caso de poseerlo) manifiesta la ausencia de mérito o demérito cuando nos autoadscribimos nuestro estado mental. Cuando estamos legitimados a acudir a una generalización existencial, tiene pleno sentido preguntarnos por las razones, hablar de convencimiento, de mérito o demérito. Ciertamente un sujeto que se autoadscribe su estado mental en perspectiva de primera persona podría acudir a una generalización existencial, el problema es que las razones que justifican tal recurso son las mismas que tiene para autoadscribirse su estado mental, a saber: ninguna; de ahí que haya de ser considerado un movimiento ilegítimo. Si el recurso a

una generalización existencial queda deslegitimado, carecemos de razones para afirmar que una autoadscripción mental puede ser capaz de eliminar la cuestión sobre la duda⁸, ya que el movimiento hacia una generalización existencial indica que la discusión sobre las razones tiene sentido, y es que aún cabría preguntarse por (y convencerse acerca de) el sujeto y el predicado de la generalización existencial. Esta situación epistémicamente desfavorable se debe a lo que el sujeto hace al autoadscribirse su estado mental: la situación (expresiva) del sujeto respecto de su supuesto objeto de conocimiento —i.e., su propio estado mental— le impide tener razones a favor o en contra de su autoadscripción. La única manera de tener razones a favor o en contra de su autoadscripción pasaría por abandonar la perspectiva de primera persona, pero esto no podría hacerse sin traicionar la idea de autoconocimiento privilegiado.

No estamos exigiendo que el sujeto deba proferir sus autoadscripciones teniendo como base razones a favor de estas. Nuestro requisito tenía que ver con el papel normativo de las razones a la hora de calificar una determinada afirmación como una de conocimiento. Decíamos que recurrir al papel expresivo para garantizar la especial seguridad podría explicar por qué no es correcto que alguien preguntase por las razones, por qué está fuera de lugar; pero que dicho movimiento esté fuera de lugar no eximía a dichas autoadscripciones de ser consideradas bajo los aspectos normativos que gobiernan nuestras atribuciones de conocimiento.

Por último, examinaremos una visión inferencialista de las afirmaciones de conocimiento que explota la posibilidad de tener conocimiento observacional no-inferencial adquirido a través de nuestra capacidad para responder a ciertos estímulos; y donde la noción de justificación presenta una dimensión

⁸La cuestión de la exclusión de la duda por necesidad lógica es ahora más diáfana. La raíz de este asunto se encuentra en la íntima conexión existente entre nuestras autoadscripciones mentales y las generalizaciones existenciales correspondientes. Si las primeras caen, las últimas también. No tiene sentido dudar de si es otro el que está en cierto estado mental, o si tengo un estado mental u otro.

inferencial, no-estática, que adquiere sentido en la consideración en conjunto del intercambio lingüístico donde las afirmaciones de conocimiento tienen lugar.

En esta clase de conocimiento no-inferencial el sujeto no hace descansar sus afirmaciones sobre el mundo en cierta evidencia, sino que estas descansan sobre ciertas capacidades del sujeto para responder con propiedad a los estímulos del entorno, a saber: aplicando el concepto adecuado ante el estímulo correspondiente. Entre los ejemplos de esta clase de conocimiento encontramos, por ejemplo, a los sexadores de pollos ya famosos en la literatura epistemológica. En Brandom (2002, 97 y ss.), Robert Brandom expone el caso de los sexadores de pollos y por qué merecen ser acreditados con conocimiento. Estos poseen conocimiento no-inferencial y digno de confianza acerca del sexo de cada pollo sin tener acceso a las características de los pollos que les hacen responder con acierto dependiendo del sexo del animal, es más, pueden estar equivocados al pensar que lo saben gracias a su aparato visual, cuando realmente su discriminación se apoya en sus capacidades olfativas. Según Brandom, los sexadores de pollos no adquieren conocimiento de manera inferencial, o identificando el sexo del animal según ciertos criterios; empero, esto no significa que sus afirmaciones no estén justificadas, dado que

[...] la justificación para esas creencias es en un sentido importante inferencial. La justificación del sujeto para creencias de este tipo depende de sacar conclusiones de una afirmación fiable antecedente. En este sentido, el sujeto que cree no está en una posición diferente a la que un observación en tercera persona estaría. (Brandom, 2002, 98) [traducción propia]

En lo que sigue, discutiremos el movimiento inferencialista que permite hacer de las respuestas del sexador a los diferentes pollos afirmaciones de conocimiento. Aunque el sexador no tuviese constancia, acceso a la evidencia que justificaría su informe perceptivo, tendría sentido preguntarse por las razones. A pesar de que el método con el que logra *saber* el sexo del pollo no descansa sobre razones

que apoyarían su afirmación, el sujeto está en posición para aportarlas, como apunta Brandom al comparar la situación del sujeto con la de otro en tercera persona.

Los sexadores de pollos podrían estar equivocados al aportar razones, podrían ofrecer razones ‘triviales’ del tipo “Simplemente porque lo creo”, “Porque tengo una conexión especial con los pollos” cuando se le preguntase cómo saben el sexo del animal; empero, los sexadores se encuentran en posición para ofrecer tales razones, mientras que el sujeto que a-expresa su estado mental no. En caso de no ofrecer las razones adecuadas, alguien podría argüir que los sexadores *no saben que* el pollo es macho (o hembra). El problema de esta objeción es que no considera que estos puedan ofrecer como razón para su afirmación de conocimiento el hecho de que han sido entrenados para responder con fiabilidad a los estímulos del entorno aplicando los conceptos adecuados⁹. Los sexadores de pollos saben cómo identificar a dichos animales según su sexo y este *saber cómo* constituye una buena razón para realizar una afirmación de conocimiento sobre el sexo del animal. Una objeción como la que hemos presentado contra la idea de que los sexadores de pollos no saben si el animal es macho o hembra puede salvarse si permitimos que el conocimiento del sexo del animal —saber que el animal es macho (o hembra)— pueda descansar sobre una capacidad práctica del sujeto, viz. saber cómo *identificar* (o *reconocer*) el sexo de los pollos. En resumen, dicha capacidad podría actuar como razón conectada inferencialmente con nuestra afirmación de conocimiento, permitiendo así evaluar dicha afirmación como una meritoria.

En cambio, el neoexpresivismo de Bar-On no podría adoptar una estrategia similar. El neoexpresivismo presentado por Bar-On no podría argüir que las autoadscripciones mentales están conectadas inferencialmente

⁹Razones como “Simplemente porque lo creo” o “Porque tengo una conexión especial con los pollos” pueden ser vistas como implicaturando conversacionalmente la proposición: “Porque he sido entrenado para responder conceptual y adecuadamente a ciertos estímulos del entorno”, es decir, lo que el hablante quiere decir con razones de ese tipo es que posee la *capacidad* para identificar el sexo del animal.

con la capacidad del sujeto para responder adecuadamente mediante vehículos lingüísticos a su condición mental. La capacidad para a-expresar autoadscriptivamente el estado mental en que nos encontramos no podría actuar como razón a favor de nuestra autoadscripción mental, permitiéndonos así evaluarla como una afirmación meritoria. No podría funcionar como una razón porque, a diferencia del caso de los sexadores, la capacidad que pone en juego el sujeto es la de a-expresar su estado mental actual, cuyo objeto no es identificar, reconocer, el estado mental que este se encuentra. La capacidad expresiva no puede actuar como una razón para la autoadscripción mental porque su función no involucra ningún proceso de convencimiento o eliminación de la duda. El papel de la a-expresión es ‘hacer visible’ a los demás el estado mental en que nos encontramos y no obtener información sobre el susodicho estado mental.

De la exposición de Brandom nos interesa recalcar los compromisos normativos que nuestras afirmaciones sobre el mundo tienen que satisfacer para ser consideradas conocimiento. Aunque Brandom se centre aquí en el conocimiento obtenido no-inferencialmente, su idea puede ser extendida a otra clase de preferencias. El carácter inferencial de la justificación permite salvar el estatus epistémico de ciertas preferencias y, lo que es más importante, logra satisfacer nuestra idea de mérito, esto es, que tenga sentido preguntar por las razones. Acudir a las relaciones inferenciales con otras afirmaciones para dar cuenta de la justificación permite dejar de lado un internismo excesivamente restrictivo. No se trata de que la afirmación de conocimiento se apoye en el acceso del sujeto a cierta evidencia a favor de su afirmación, sino de que tenga sentido que el sujeto pueda acudir a otra afirmación que funcione como una razón conectada inferencialmente a la afirmación principal.

El problema con las autoadcripciones mentales, entendidas según el modelo neoexpresivista de Bar-On, no es que su función expresiva las caracterice como preferencias realizadas sin posesión de evidencia o razones a su favor, sino que su carácter expresivo incapacita al sujeto para *aportar* razones a favor o en contra de su autoadscripción mental. La ausencia de razones

[*groundlessness*] propia del fenómeno de la autoadscripción mental caracteriza la situación del sujeto dentro de lo que otros han llamado “el juego de dar y pedir razones”. Tal y como pretendíamos hacer ver con la última cita de Wittgenstein, no puede decirse con sentido del sujeto que se encuentre en una posición adecuada para ofrecer razones a no ser que abandone su perspectiva de primera persona. No obstante, abandonar esta posición privilegiada nos obligaría a caracterizar al autoconocimiento desde una perspectiva de tercera persona, haciendo flaco favor a su carácter privilegiado.

En suma, a lo largo de toda esta discusión no hemos intentado negar la posibilidad de autoconocimiento en general, sino más bien que nuestras autoadscripciones mentales, entendidas bajo el neoexpresivismo de Bar-On, articulasen cierta clase de conocimiento especial en perspectiva de primera persona. Convenimos con Bar-On que nuestros actos autoadscriptivos en virtud de los cuales nos atribuimos un estado mental de forma ‘espontánea’, sin participar de un proceso inferencial, no tienen por función dotar al sujeto con conocimiento acerca de su estado mental presente, pero discrepamos en el hecho de que nuestras autoadscripciones puedan ser consideradas afirmaciones de autoconocimiento a la luz de su neoexpresivismo.

5. CONCLUSIONES

A lo largo de este trabajo, nuestro objetivo principal ha sido mostrar que las autoadscripciones mentales tal y como son concebidas por el neoexpresivismo de Bar-On no pueden ser consideradas afirmaciones de autoconocimiento. Nuestra conclusión descansa sobre la tesis de que lo esencial en las atribuciones de conocimiento (y autoconocimiento) no es describir una relación cognitiva, viz. creencia, que el sujeto mantiene con una proposición verdadera, sino atribuir mérito o demérito al sujeto teniendo en cuenta la acción sobre la que descansa su afirmación de conocimiento.

En la segunda sección, caracterizamos el mérito epistémico atendiendo a la presencia de razones a las que el sujeto pudiera acudir en caso de buscarla justificar su afirmación inicial, lo cual distinguimos claramente de la formación de juicios o creencias sobre la base de razones o evidencia. Precisamente nuestras autoadscripciones mentales son proferidas sin fundamentos o razones —son *groundless*—; empero, esto no las exime de poder ser evaluadas conforme a sus conexiones inferenciales con otras afirmaciones, cuya función consistiría en eliminar cualquier resquicio de duda acerca de la condición mental del sujeto.

Tras presentar la postura de Bar-On, según la cual nuestras autoadscripciones mentales son *expresiones* lingüísticas y autoadscriptivas de nuestros propios estados mentales, identificamos el problema de la explicación de Bar-On. El carácter expresivo de las autoadscripciones mentales no podía ser conjugado con la dimensión epistémica, dado que las autoadscripciones mentales no podían ser evaluadas conforme al mérito o demérito: no tenía sentido hablar de duda en dichas situaciones, ni el sujeto podía conectar su autoadscripción mental con otras afirmaciones que funcionasen como razones. En definitiva, a lo largo de nuestra evaluación, hemos concedido a Bar-On el éxito de su alternativa expresivista a la hora de explicar las peculiaridades del privilegio de la primera persona, hasta tal punto que, *pace* Bar-On, su propia caracterización de la perspectiva de la primera persona ha terminado por excluir a las autoadscripciones mentales del ámbito de lo epistémico.

REFERENCIAS

- ANSCOMBE, G. E. M. (1957/1991). *Intention*. Oxford: Basil Blackwell.
Citado aquí por la traducción castellana: *Intención*. Barcelona: Paidós.
- BAR-ON, D. (2004). *Speaking My Mind: Expression and Self-knowledge*. Oxford: Oxford University Press.
- (2009). “First-Person Authority: Dualism, Constitutivism, and Neo-Expressivism”. *Erkenntnis*, 71 (1): 53-71.

- (2011). “Neo-Expressivism: Avowals’ Security and Privileged Self-Knowledge”. En A. Hatzimoysis (Ed.), *Self-knowledge* (189-201), Oxford: Oxford University Press.
- (2015). “Transparency, expression, and self-knowledge”. *Philosophical Explorations*, 18(2): 134-152.
- BAR-ON, D. y CHRISMAN, M. (2009). “Ethical Neo-Expressivism”. En R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 4; 133-165). Oxford: Oxford University Press.
- BAR-ON, D. y SIAS, J. (2013). “Varieties of Expressivism”. *Philosophy Compass*, 8(8): 699-713.
- BOYLE, M. (2009). “Two kinds of self-knowledge”. *Philosophy and Phenomenological Research*, 78(1): 133-164.
- BRANDON, R. (2002). “Non-inferential knowledge, perceptual experience, and secondary qualities: Placing McDowell’s empiricism”. En N. H. Smith (Ed.), *Reading McDowell. On Mind and World* (92-105). New York: Routledge.
- CHRISMAN, M. (2009). “Expressivism, Truth, and (Self-)Knowledge”. *Philosophers’ Imprint*, 9(3): 1-26.
- (2010). “From Epistemic Expressivism to Epistemic Inferentialism”. En A. Haddock, A. Millar y D. Pritchard (Eds.), *Social Epistemology* (112-128). Oxford: Oxford University Press.
- (2012). “Epistemic Expressivism”. *Philosophy Compass*, 7(2): 118–126.
- MORAN, R. (2001). *Authority and Estrangement: an Essay on Self-knowledge*. Princeton (NJ): Princeton University Press.
- RÖDL, S. (2007). *Self-consciousness*. Cambridge (MA): Harvard University Press.
- RYLE, G. (1949/2009). *The Concept of Mind*. London: Hutchinson. Citado aquí por *The Concept of Mind*, London: Routledge.
- WITTGENSTEIN, L. (1953/2010). *Philosophical Investigations* (G.E.M. Anscombe, trad.). Oxford: Basis Blackwell. Citado aquí por la traducción castellana: *Investigaciones filosóficas*. Barcelona: Crítica.

——(1958/2007). *Preliminary Studies for the “Philosophical Investigations”, Generally known as The Blue and Brown Books*. Oxford: Basil Blackwell.
Citado aquí por la traducción castellana: *Los cuadernos azul y marrón* (4^a ed.). Madrid: Tecnos.

WRIGHT, C. (1998). “Self-knowledge: the Wittgensteinian Legacy”. *Royal Institute of Philosophy Supplement*, 43: 101-122.

——(2001). “The Problem of Self-knowledge (I)”. En *Rails to Infinity: Essays on Themes from Wittgenstein’s Philosophical Investigations* (319-344). Cambridge (MA): Harvard University Press.